Adaptive Feedback Cancellation in Hearing Aids using a Sinusoidal Near-End Signal Model

Ngo, Kim; van Waterschoot, Toon; Christensen, Mads Græsbøll; Moonen, Marc; Jensen, Søren Holdt; Wouters, Jan

Published in:
I E E E International Conference on Acoustics, Speech and Signal Processing. Proceedings

DOI (link to publication from Publisher):
10.1109/ICASSP.2010.5496063

Publication date:
2010

Document Version
Accepteret manuscript, peer-review version

ADAPTIVE FEEDBACK CANCELLATION IN HEARING AIDS USING A SINUSOIDAL NEAR-END SIGNAL MODEL

Kim Ngo[1], *Toon van Waterschoot*[1], *Mads Græsbøll Christensen*[2], *Marc Moonen*[1], *Søren Holdt Jensen*[2] *and Jan Wouters*[3]

[1]Katholieke Universiteit Leuven, ESAT-SCD, Kasteelpark Arenberg 10, B-3001 Leuven, Belgium
[2]Aalborg University, Dept. Electronic Systems, Niels Jernes Vej 12, DK-9220 Aalborg, Denmark
[3]Katholieke Universiteit Leuven, ExpORL, O. & N2, Herestraat 49/721, B-3000 Leuven, Belgium

ABSTRACT

Acoustic feedback is a well-known problem in hearing aids, which is caused by the undesired acoustic coupling between the loudspeaker and the microphone. Acoustic feedback limits the maximum amplification that can be used in the hearing aid without making it unstable. The goal of adaptive feedback cancellation (AFC) is to adaptively model the feedback path and estimate the feedback signal, which is then subtracted from the microphone signal. The main problem in identifying the feedback path model is the correlation between the near-end signal and the loudspeaker signal, which is caused by the closed signal loop. A possible solution to this problem is to use the prediction error method (PEM)-based AFC with a linear prediction (LP) model for the near-end signal. In this paper, a modification to the PEM-based AFC is presented where the LP model is replaced by a sinusoidal near-end signal model. More specifically, it is shown that using frequency estimation techniques to estimate the sinusoidal near-end signal model improves the performance of the PEM-based AFC compared to using a LP model. Simulation results for a hearing aid scenario indicate a significant improvement in terms of misadjustment and maximum stable gain increase.

Index Terms— Adaptive Feedback Cancellation, Frequency Estimation, Decorrelation, Hearing Aids.

1. INTRODUCTION

Acoustic feedback is a well-known problem in hearing aids, which is caused by the undesired acoustic coupling between the loudspeaker and the microphone. Acoustic feedback limits the maximum amplification that can be used in a hearing aid if howling, due to instability, is to be avoided. In many cases this maximum amplification is too small to compensate for the hearing loss, which makes feedback cancellation algorithms an important component in hearing aids.

The goal of adaptive feedback cancellation (AFC) is to adaptively model the feedback path and estimate the feedback signal, which is then subtracted from the microphone signal. The main problem in identifying the feedback path model is the correlation between the near-end signal and the loudspeaker signal, which is caused by the closed signal loop. This correlation problem causes standard adaptive filtering algorithms to converge to a biased solution. The challenge is therefore to reduce the correlation between the near-end signal and the loudspeaker signal. Typically, there exist two approaches to this decorrelation [1], i.e., decorrelation in the closed signal loop and decorrelation in the adaptive filtering circuit. Recently proposed methods for decorrelation in the closed signal loop consist in the insertion of all-pass filters [2] in the forward path of the hearing aid or in clipping [3] of the feedback signal arriving at the microphone. Alternatively, an unbiased identification of the feedback path model can be achieved by applying decorrelation in the adaptive filtering circuit, i.e., by first prefiltering the loudspeaker and microphone signals with the inverse near-end signal model before feeding these signals to the adaptive filtering algorithm [4], [5]. The near-end signal model and the feedback path model can be jointly estimated using the so-called prediction error method (PEM). For near-end speech signals, a linear prediction (LP) model is commonly used in hearing aids [4]. For audio signals a cascade of a constrained pole-zero LP (CPZLP) model with a LP model has been proposed [5].

In this paper, the goal is to use a sinusoidal model for the near-end signal instead of a LP model in PEM-based AFC. The sinusoidal near-end signal model can be fitted into the prediction error framework by exploiting LP properties of sinusoidal signals [6]. In [7] a frequency estimation method is proposed that is based on CPZLP, which is used as the near-end signal model. The frequencies are then suppressed by using notch filters implemented as second-order pole-zero filters. In this paper, the CPZLP is replaced by fundamental frequency estimation methods based on subspace shift-invariance and subspace orthogonality, and optimal filtering [8]. The sinusoidal components are then suppressed by a cascade of notch filters centered at the frequencies of the sinusoidal components that are here assumed to be integer multiples of a fundamental frequency. The different PEM-based AFC algorithms are compared using speech signals in a hearing aid configuration. The AFC performance is evaluated in terms of maximum stable gain (MSG), misadjustment and sound quality.

The paper is organized as follows. Section 2 describes the adaptive feedback cancellation concept. In section 3, the concept of using a sinusoidal near-end signal model is explained. Section 4 describes the different frequency estimation methods used. In Section 5, simulation results are presented. The work is summarized in Section 6.

2. ADAPTIVE FEEDBACK CANCELLATION

The adaptive feedback cancellation concept is shown in Fig. 1. The microphone signal is given by

$$y(t) = v(t) + x(t) = v(t) + F(q,t)u(t) \tag{1}$$

This research work was carried out at the ESAT laboratory of the Katholieke Universiteit Leuven, in the frame of the Marie-Curie Fellowship EST-SIGNAL program (http://est-signal.i3s.unice.fr) under contract No. MEST-CT-2005-021175, and the Concerted Research Action GOA-AMBioRICS and the Belgian Programme on Interuniversity Attraction Poles initiated by the Belgian Federal Science Policy Office IUAP P6/04 (DYSCO, 'Dynamical systems, control and optimization', 2007-2011). The scientific responsibility is assumed by its authors.

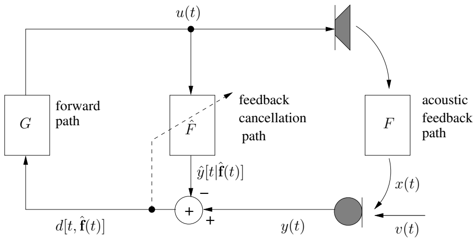


Fig. 1. Adaptive feedback cancellation (AFC).

where q denotes the time shift operator and t is the dicrete time variable. $F(q,t)$ is the feedback path between the loudspeaker and the microphone, $v(t)$ is the near-end signal, $x(t)$ is the feedback signal. The forward path $G(q,t)$ maps the microphone signal $y(t)$, possibly after AFC, to the loudspeaker signal $u(t)$. The concept of the AFC is to place an estimated finite impulse response (FIR) adaptive filter $\hat{F}$ in parallel with the feedback path, having the loudspeaker signal as input and microphone signal as the desired output. The feedback canceller $\hat{F}$ produces an estimate of the feedback signal $x(t)$ which is then subtracted from the microphone signal $y(t)$. The feedback-compensated signal is given by

$$d(t) = v(t) + [F(q,t) - \hat{F}(q,t)]u(t). \tag{2}$$

The main problem in identifying the feedback path model is the correlation between the near-end signal and the loudspeaker signal, which causes standard adaptive filtering algorithms to converge to a biased solution. This means that the adaptive filter does not only predict and cancel the feedback component in the microphone signal, but also part of the near-end signal, which results in a distorted feedback-compensated signal $d(t)$. Alternatively, an unbiased identification of the feedback path model can be achieved by applying decorrelation in the adaptive filtering circuit, i.e., by first prefiltering the loudspeaker and microphone signals with the inverse near-end signal model before feeding these signals to the adaptive filtering algorithm. The near-end signal model and the feedback path model can be jointly estimated using the so-called prediction error method (PEM). For details on the PEM-based AFC we refer to [1], [4], [5].

3. SINUSOIDAL NEAR-END SIGNAL MODEL

The near-end signal $v(t)$ and hence the feedback-compensated signal $d(t)$ are assumed to consist of a sum of real sinusoids and additive noise,

$$d(t) = \sum_{n=1}^{P} A_n \cos(\omega_n t + \phi_n) + r(t), \quad t = 1, ..., M \tag{3}$$

with A_n the amplitude, $\omega_n \in [0, \pi]$ the radial frequency, and $\phi_n \in [0, 2\pi)$ the phase of the *n*th sinusoid, and $r(t)$ the noise.

In this paper, the goal is to use a sinusoidal model of the near-end signal instead of a LP model in PEM-based AFC. A particular class of parametric methods exploits the LP property of sinusoidal signals. It is well known that a sum of *P* sinusoids can be described exactly using an all-pole model of order 2*P*, with mirror symmetric LP coefficients. However, it has been shown that the all-pole model is not exact when noise is added, and in this case a pole-zero model of order 2*P* should be used [6]. Still, by constraining the poles and zeros to lie on common radial lines in the *z*-plane, the number of unknown parameters in the pole-zero model can be limited to *P* and the LP parameters can be uniquely related to the unknown frequencies [7]. The CPZLP model can be written as

$$d(t) = \left(\prod_{n=1}^{P} \frac{1 - 2\rho \cos\omega_n z^{-1} + \rho^2 z^{-2}}{1 - 2\cos\omega_n z^{-1} + z^{-2}} \right) e(t) \tag{4}$$

where ω_n denotes the frequencies and ρ the pole radius.

In case of colored noise in the sinusoidal near-end signal model, an additional prediction error filter can be cascaded with the CPZLP model. The former then predicts the noise components and the latter predicts the sinusoidal components in the near-end signal [5]. In this paper, a CPZLP model is used for the sinusoidal components and for the noise components a conventional all-pole model is chosen.

In [7] a frequency estimation method is proposed that is based on the CPZLP model, and applied to PEM-based AFC in [5]. In this paper, the CPZLP frequency estimation method is replaced by fundamental frequency estimation methods based on subspace shift-invariance and subspace orthogonality, and optimal filtering as described in [8]. The sinusoidal components are then suppressed by a cascade of notch filters centered at the frequencies of the sinusoidal components that are here assumed to be integer multiples of a fundamental frequency.

4. SINUSOIDAL FREQUENCY ESTIMATION

In this section, different methods to estimate the sinusoidal frequencies are briefly introduced and further details can be found in [7] [8]. In several of the methods, namely those based on pitch estimation [8], it is assumed that the sinusoids are having frequencies that are integer multiples of a fundamental frequency ω_0, i.e., $\omega_n = \omega_0 n$. This follows naturally from voiced speech being quasi-periodic. This assumption is not made in the CPZLP method where all the frequencies are estimated independently.

4.1. CPZLP based frequency estimation

The CPZLP minimization criterion is given by

$$\min_{\boldsymbol{\omega}} V(\boldsymbol{\omega}) = \min_{\boldsymbol{\omega}} \frac{1}{M} \sum_{t=1}^{M} e^2(t, \boldsymbol{\omega}) \tag{5}$$

with the residual signal defined as the output from the prediction error filter

$$e(t, \boldsymbol{\omega}) = \left(\prod_{n=1}^{P} \frac{1 - 2\cos\omega_n z^{-1} + z^{-2}}{1 - 2\rho \cos\omega_n z^{-1} + \rho^2 z^{-2}} \right) d(t) \tag{6}$$

and $\boldsymbol{\omega} = [\omega_1 \quad ... \quad \omega_P]^T$. The CPZLP minimization in (5)-(6) can be solved in a decoupled fashion, using an iterative line search optimization [7].

4.2. Subspace-orthogonality-based pitch estimation

The idea behind subspace methods is to divide the full space into a signal subspace containing the signal of interest and its orthogonal complement, the noise subspace. The subspace orthogonality method is based on the observation that the sinusoids in (3) are all orthogonal to the noise subspace. The covariance matrix of the observed signal in (3) can be shown to be

$$\mathbf{R} = \mathrm{E}\{\tilde{\mathbf{d}}(t)\tilde{\mathbf{d}}^H(t)\} \tag{7}$$
$$= \mathbf{ZPZ}^H + \sigma^2 \mathbf{I} \tag{8}$$

where $(\cdot)^H$ denotes Hermitian transpose and $\tilde{\mathbf{d}}(t)$ is a vector containing M consecutive samples of the analytical counterpart of the feedback-compensated signal $d(t)$ [8]. Furthermore, $\mathbf{Z}$ is a Vandermonde matrix containing the sinusoids of the model in (3), and $\mathbf{P}$ is the covariance matrix of the amplitudes, which can be shown to be diagonal under certain conditions. Finally, σ^2 denotes the variance of the additive noise, and $\mathbf{I}$ is the identity matrix. In the presence of colored noise, it is required that pre-whitening is applied, as the model in (8) would otherwise be invalid. Exploiting the fact that the noise subspace eigenvectors $\mathbf{G}$ are orthogonal to the columns of the matrix $\mathbf{Z}$, it follows that the the fundamental frequency ω_0 can be estimated as

$$\hat{\omega}_0 = \arg\min_{\omega_0} ||\mathbf{Z}^H\mathbf{G}||_F^2, \tag{9}$$

where $\mathbf{Z}$ depends on ω_0. More specifically, the matrix $\mathbf{G}$ is constructed from the $M-2P$ least significant eigenvectors of $\mathbf{R}$.

4.3. Subspace-shift-invariance based pitch estimation

The next method is based on a particular property of the signal subspace generated by signals as in (3), namely the shift-invariance property. The signal subspace is spanned by the columns of the matrix $\mathbf{S}$ formed from the $2P$ most significant eigenvectors of $\mathbf{R}$. Two matrices $\underline{\mathbf{S}}$ and $\overline{\mathbf{S}}$ are constructed by removing the last and first row of the matrix $\mathbf{S}$ which can be shown to be related by a linear transform as $\overline{\mathbf{S}} = \underline{\mathbf{S}}\mathbf{\Xi}$. The problem of finding the fundamental frequency can then be seen as a fitting problem, i.e.

$$\overline{\mathbf{S}} \approx \underline{\mathbf{S}}\mathbf{Q}\tilde{\mathbf{D}}\mathbf{Q}^{-1} \tag{10}$$

where $\tilde{\mathbf{D}} = \text{diag}\big([e^{j\omega} \ldots \;\; e^{j\omega 2P}]\big)$ is a diagonal matrix containing the unknown fundamental frequency. The matrix $\mathbf{Q}$ contains the eigenvectors of the matrix $\hat{\mathbf{\Xi}} = (\underline{\mathbf{S}}^H\underline{\mathbf{S}})^{-1}\underline{\mathbf{S}}^H\overline{\mathbf{S}}$. The fundamental frequency can then be estimated as

$$\hat{\omega}_0 = \arg\min_{\omega_0} ||\overline{\mathbf{S}} - \underline{\mathbf{S}}\mathbf{Q}\tilde{\mathbf{D}}\mathbf{Q}^{-1}||_F^2, \tag{11}$$

which can be simplified significantly, as shown in [8].

4.4. Optimal-filtering-based pitch estimation

The final estimator is based on filtering of the feedback-compensated signal. The idea behind pitch estimation based on filtering is to find a set of filters that pass power undistorted at the harmonic frequencies $\omega_0 n$, while minimizing the power at all other frequencies. This filter design problem can be stated mathematically as

$$\min_{\mathbf{h}} \mathbf{h}^H\mathbf{R}\mathbf{h} \quad \text{s.t.} \quad \mathbf{h}^H\mathbf{z}(\omega_0 n) = 1, \quad \text{for} \quad n = 1, ..., P, \tag{12}$$

where $\mathbf{h}^H$ is the length M impulse response of the filter and $\mathbf{z}(\omega) = [e^{-j\omega 0} \; \ldots \; e^{-j\omega(M-1)}]$. Using the Lagrange multiplier method, the optimal filters can be shown to be

$$\mathbf{h} = \mathbf{R}^{-1}\mathbf{Z}\big(\mathbf{Z}^H\mathbf{R}^{-1}\mathbf{Z}\big)\mathbf{1} \tag{13}$$

with $\mathbf{1} = [1 \;\; \ldots \;\; 1]^T$. This filter is signal adaptive and depends on the unknown fundamental frequency. Intuitively, one can obtain a fundamental frequency estimate by filtering the signal using the optimal filters for various fundamental frequencies and then picking the one for which the output power is maximized, i.e.,

$$\hat{\omega}_0 = \arg\max_{\omega_0} \mathbf{1}^H\big(\mathbf{Z}^H\mathbf{R}^{-1}\mathbf{Z}\big)^{-1}\mathbf{1}. \tag{14}$$

This method has demonstrated to have a number of desirable features, namely excellent statistical performance and robustness towards periodic interference [8].

5. EVALUATION

Simulation results are presented in which different frequency estimation methods, namely CPZLP, subspace and optimal filtering methods, are compared in a PEM-based AFC approach with cascaded near-end signal models in a hearing aid setup. The near-end sinusoidal model order is set to P = 15 and the near-end noise model order is set to 30. Both near-end signal models are estimated using 50% overlapping data windows of length M = 320 samples. The NLMS adaptive filter length is set equal to the acoustic feedback path length, i.e., n_F = 200. The near-end signal is a 30 s speech signal at f_s= 16 kHz. The forward path gain $K(t)$ is set 3 dB below the maximum stable gain (MSG) without feedback cancellation.

To assess the performance of the AFC algorithm the following measures are used. The achievable amplification before instability occurs is measured by the MSG, which is defined as

$$\text{MSG}(t) = -20\log_{10}\left[\max_{\omega\in\mathcal{P}} |J(\omega,t)[F(\omega,t) - \hat{F}(\omega,t)]|\right] \tag{15}$$

where $J(q,t) = \frac{G(q,t)}{K(t)}$ denotes the forward path transfer function without the amplification gain $K(t)$, and $\mathcal{P}$ denotes the set of frequencies at with the feedback signal $x(t)$ is in phase with the near-end signal $v(t)$. The misadjustment between the estimated feedback path $\hat{\mathbf{f}}(t)$ and the true feedback path $\mathbf{f}$ represents the accuracy of the feedback path estimation and is defined as,

$$\text{MA}_F = 20\log_{10}\frac{||\hat{\mathbf{f}}(t) - \mathbf{f}||_2}{||\mathbf{f}||_2}. \tag{16}$$

A frequency-weighted log-spectral signal distortion (SD) is used to measure the sound quality, defined as

$$\text{SD}(t) = \sqrt{\int_0^{fs/2} w_{\text{ERB}}(f)\left(10\log_{10}\frac{S_d(f,t)}{S_v(f,t)}\right)^2 df} \tag{17}$$

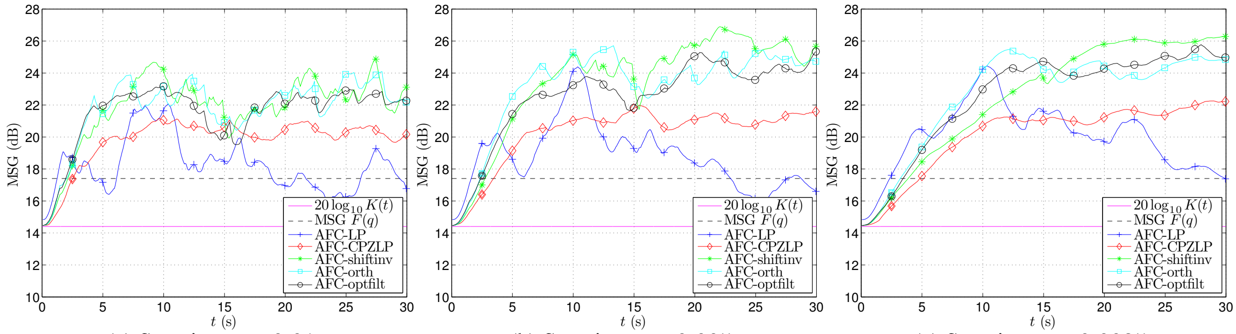
where $S_d(f,t)$ and $S_v(f,t)$ denote the short-term PSD of the feedback-compensated signal and the near-end signal, respectively, and $w_{\text{ERB}}(f)$ is a frequency-weighting factor giving equal weight for each auditory critical band [9]. The integration in (17) is approximated by a summation over the DFT frequency bins and the mean value of the SD measure is used in the evaluation.

5.1. Simulation results

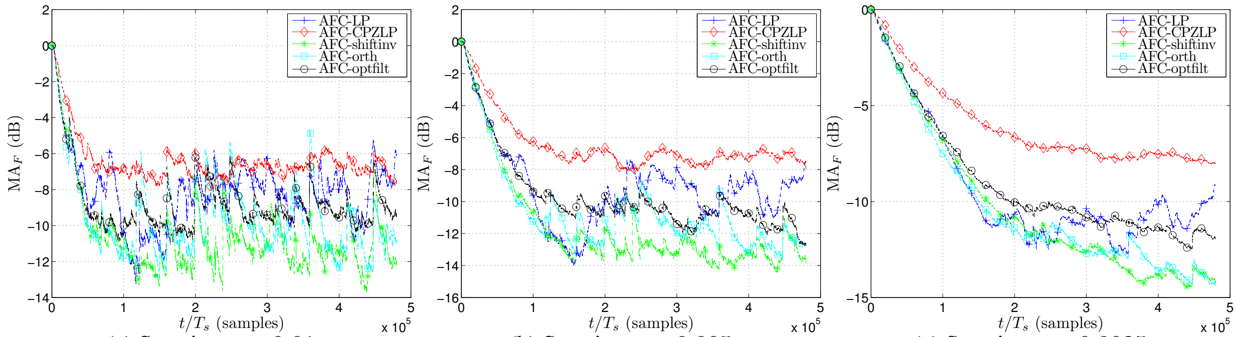
The instantaneous value of the MSG(t) is shown in Fig. 2 for different stepsize μ and the corresponding misadjustment is shown in Fig. 3. The MSG(t) curves have been smoothed with a one-pole low-pass filter to improve the clarity of the figures. The instantaneous value of the forward path gain $20\log_{10} K(t)$ and the MSG without acoustic feedback control (MSG $F(q)$) are also shown.

The AFC-LP is included as a reference since a single all-pole model is currently used in PEM-based AFC in hearing aids [4]. At some point the MSG in the AFC-LP decreases and even gets close to instability. Compared to the AFC-CPZLP, the MSG in this case seems to be more stable with an overall higher MSG compared to the AFC-LP even though the mistadjustment is lower for AFC-LP. The benefit of AFC-CPZLP can be explained by the benefit of using a cascaded near-end signal model. A cascade of near-end signal models removes the coloring and periodicity (due to glottal excitation) in voiced speech segments. On the other hand, a single short-term predictor fails to remove the periodicity, which causes the loudspeaker signal still being correlated with the near-end signal during voiced speech.

The MSG is in general higher using AFC-shiftinv, AFC-orth and AFC-optfilt compared to the exisiting methods AFC-LP and AFC-CPZLP, which supports the conjecture that an accurate estimation of the near-end signal model results in a better decorrelation and hence an increase in MSG. Using lower stepsize shows a significantly better convergence behavior for AFC-shiftinv, AFC-orth and AFC-optfilt compared to AFC-CPZLP. From these results, it is clear that the frequency estimation methods have a great impact on the AFC performance. On the other hand, it is worth noting that the



(a) Stepsize $\mu = 0.01$ (b) Stepsize $\mu = 0.005$ (c) Stepsize $\mu = 0.0025$



(a) Stepsize $\mu = 0.01$ (b) Stepsize $\mu = 0.005$ (c) Stepsize $\mu = 0.0025$

Fig. 3. Misadjustment between the estimated feedback path $\hat{\mathbf{f}}(t)$ and the true feedback path $\mathbf{f}$.

Table 1. Sound quality

	Mean (SD) [dB]		
Method	μ=0.01	μ=0.005	μ=0.0025
LP	2.3965	2.1486	2.1374
CPZLP	4.2801	4.2107	4.4317
Shiftinv	2.7654	2.6536	3.0365
Orth	3.2171	3.0276	3.2341
Optfilt	3.5041	3.3007	3.5540

choice of the stepsize seems to have a great impact on the convergence for AFC-shiftinv, AFC-orth and AFC-optfilt, whereas AFC-CPZLP seems to stabilize faster but at a larger error.

The sound quality in terms of distortion is shown Table 1, and amongst the PEM-based AFC algorithms, the AFC-shiftinv yields the lowest SD while still maintaining a MSG value comparable to AFC-orth and AFC-optfilt. The AFC-LP algortihm provides the best sound quality but this comes at the cost of poor MSG. In terms of sound quality, the SD measure shows that the distortion is highest when the CPZLP method is used.

6. CONCLUSION

In this paper, a sinusoidal near-end signal model is introduced instead of a linear prediction model typically used in PEM-based AFC. Furthermore, different frequency estimation methods in PEM-based AFC have been evaluated and compared in terms of achievable amplification, sound quality and misadjustment of the estimated feedback path. It is shown, that the performance of a PEM-based AFC with cascaded near-end signal models can be further improved by using pitch estimation methods where the sinusoidal frequencies are an integer multiple of a fundamental frequency, which is different compared CPZLP where all frequencies are estimated. The pitch estimation methods considered here are based on subspace and optimal filtering. Overall the achievable amplification in terms of MSG is higher and the misadjustment is lower using subspace and optimal filtering methods. Since the sinusoidal near-end signal model cascaded with an all-pole model is able to whiten the near-end signal component in the microphone signal more effectively, a significant AFC performance improvement is obtained.

7. REFERENCES

[1] T. van Waterschoot and M. Moonen, “50 years of acoustic feedback control: state of the art and future challenges,” *Proc. IEEE*, submitted for publication, Feb. 2009, ESAT-SISTA Technical Report TR 08-13, Katholieke Universiteit Leuven, Belgium.

[2] C. Boukis, D. P. Mandic, and A. G. Constantinides, “Toward bias minimization in acoustic feedback cancellation systems,” *J. Acoust. Soc. Am.*, vol. 121, no. 3, pp. 1529–1537, Mar. 2007.

[3] D. J. Freed, “Adaptive feedback cancellation in hearing aids with clipping in the feedback path,” *J. Acoust. Soc. Am.*, vol. 123, no. 3, pp. 1618–1626, Mar. 2008.

[4] A. Spriet, I. Proudler, M. Moonen, and J. Wouters, “Adaptive feedback cancellation in hearing aids with linear prediction of the desired signal,” *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3749–3763, Oct. 2005.

[5] T. van Waterschoot and M. Moonen, “Adaptive feedback cancellation for audio applications,” *Signal Processing*, vol. 89, no. 11, pp. 2185–2201, Nov. 2009.

[6] Y. T. Chan, J. M M. Lavoie, and J. B. Plant, “A parameter estimation approach to estimation of frequencies of sinusoids,” *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 29, no. 2, pp. 214–219, Apr. 1981.

[7] T. van Waterschoot and M. Moonen, “Constrained pole-zero linear prediction: an efficient and near-optimal method for multitone frequency estimation,” in *Proc. 16th European Signal Process. Conf. (EUSIPCO ’08)*, Lausanne, Switzerland, Aug. 2008.

[8] M. G. Christensen and A. Jakobsson, *Multi-Pitch Estimation*, Morgan & Claypool, 2009.

[9] Moore. B, *An Introduction to the Psychology of Hearing*, Academic Press, 5th ed edition, 2003.